

WIRTSCHAFTSINFORMATIK 2

PANDAS DATAFRAMES TEIL 2

PROF. DR. BERND BLÜMEL, PROF. DR. CHRISTIAN
BOCKERMANN, PROF. DR. VOLKER KLINGSPOR

HOCHSCHULE BOCHUM

WINTERSEMESTER 2024/2025

Was Sie letzte Woche gelernt haben sollten:

- Zugriff auf Zeilen/Spalten/Zellen über die Position mit `.iloc`
- Filtern von DataFrames (Zeilen)
- Aggregation von Zeilen bzw. Spalten mit Aggregationsfunktionen

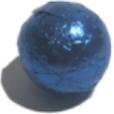
Beispiel: Kundendaten vom Schoko-Shop



Warenkorb: 0 Artikel, 0,00 €

Kategorien

Schokolade

| | | | |
|--|---|--|--|
| <p>Rübezahl Schokolade</p>  <p>Schokokugel in gold</p> <p>Preis: 0,05 € 🛒</p> | <p>Rübezahl Schokolade</p>  <p>Schokokugel in rot</p> <p>Preis: 0,05 € 🛒</p> | <p>Rübezahl Schokolade</p>  <p>Schokokugel in blau</p> <p>Artikel aktuell leider nicht verfügbar.</p> | <p>Rübezahl Schokolade</p>  <p>Schokokugel in grün</p> <p>Preis: 0,05 € 🛒</p> |
|--|---|--|--|

Shop-Daten

shop-hsbo.de
Version: 0.0.1

Filtern und Aggregieren

Die Tabelle zeigt einen Ausschnitt aus Kundeninformationen vom Schoko-Shop vom Bochumer Familientag:

| KundeNr | Kundengruppe | Altersgruppe | Artikel | Betrag |
|---------|--------------|--------------|---------|--------|
| 282 | Schulkind | 6-10 | 4 | 0.15 |
| 269 | Mitarbeiter | 41-50 | 2 | 0.1 |
| 261 | KigaKind | 0-5 | 2 | 0.05 |
| 280 | Schulkind | 6-10 | 3 | 0.1 |

```
import pandas as pd

kd = pd.read_csv('https://data.hsbo.de/kunden.csv')
```

Filtern und Aggregieren

| KundeNr | Kundengruppe | Altersgruppe | Artikel | Betrag |
|---------|--------------|--------------|---------|--------|
| 282 | Schulkind | 6-10 | 4 | 0.15 |
| 269 | Mitarbeiter | 41-50 | 2 | 0.1 |
| 261 | KigaKind | 0-5 | 2 | 0.05 |
| 280 | Schulkind | 6-10 | 3 | 0.1 |

- Was ist der Gesamtumsatz der Kundengruppe **KigaKind**?
- Wieviele Artikel wurden insgesamt in der Altersgruppe **6-10** bestellt?
- Wieviele Kunden gibt es in der Gruppe **Mitarbeiter**?

Heutige Übung

- Übungsaufgaben mit Hilfe von IBIX
- Bonuspunkt-Aufgaben **ab Donnerstag** verfügbar

Heutige Übung

- Übungsaufgaben mit Hilfe von IBIX
- Bonuspunkt-Aufgaben **ab Donnerstag** verfügbar

Selbststudium nächste Woche

- Thema: *Daten Gruppieren (Pandas DataFrames)*
- Foliensatz *Pandas Group-By* (datascience.hs-bochum.de)
- Fragen gerne bis Dienstag per Mail an den Dozenten

Wochenaufgabe

- Verfügbar auf

<https://datascience.hs-bochum.de>
(Vorlesungen -> Semester -> Wirtschaftsinformatik 2)